

BAB VI

PENUTUP

6.1. Kesimpulan

Pada eksperimen yang dilakukan menggunakan *dataset* diambil dari komentar video klip Raisa – kali kedua pada tanggal 2 Agustus 2021 terdiri dari 6480 (enam ribu empat ratus delapan puluh) dan lima kolom. Dari hasil evaluasi terhadap algoritma dan pemaknaan pada setiap *cluster* yang telah divalidasi oleh *expert judgement* yang dilakukan oleh Ike Devi Sulistyaningtyas S.Sos., M.Si. dapat menjawab tujuan dari eksperimen ini. Maka hasil jawaban sebagai berikut.

1. Implementasi *clustering* menggunakan metode *K-means*. Hasil evaluasi untuk mengukur kualitas *clustering* menggunakan nilai *silhouette score* dengan *metric Eucladien distance*. Hasil dari nilai *silhouette score* yang didapatkan menunjukkan cukup baik.
2. Implementasi *word embedding* menggunakan metode *Word2vec*. Dapat dilihat dari hasil nilai *silhouette score* yang menunjukkan cukup baik sehingga proses *clustering* data teks menggunakan *word embedding* metode *Word2vec* dapat dilakukan dengan baik.
3. Hasil pemaknaan pada setiap *cluster* dengan pendekatan segitiga semantik Ogden dan Richard's dapat dimaknai dengan baik sehingga dapat menghasilkan pengetahuan. Dari pemaknaan tersebut diketahui bahwa audiens tidak hanya menyukai lagu namun juga menyukai aktor dan aktris pada video klip tersebut.

Maka dapat disimpulkan penggunaan model yang dibuat hingga pemaknaan cukup baik dan dapat digunakan pada *dataset* teks yang lain. Selain itu pemaknaan dari hasil clustering dapat digunakan pada penelitian selanjutnya jika diperlukan.

6.2. Saran

Dari hasil yang dihasilkan, maka saran sebagai pertimbangan untuk penelitian yang akan dilakukan pada masa mendatang sebagai berikut.

1. Eksperimen ini hanya menggunakan *dataset* yang diambil dari YouTube sehingga untuk kedepannya, diharap mencoba menggunakan *dataset* teks dari *social media* yang lain atau dari dokumen – dokumen yang menggunakan data teks.
2. Eksperimen ini menggunakan metode *K-means* pada proses *clustering* sehingga diharapkan pada penelitian berikutnya menggunakan metode yang berbeda atau menggunakan metode yang sama dengan menggabungkan teknik lain yang mungkin dapat menjadi pembanding.
3. Pada proses *word embedding* penulis menggunakan metode *Word2vec* sehingga pada penelitian yang akan datang, diharapkan menggunakan metode yang lain atau mengoptimalkan metode tersebut.

Dengan adanya saran seperti di atas, diharapkan untuk penelitian yang akan data menjadi lebih baik dan lebih bermanfaat.

DAFTAR PUSTAKA

- [1] T. Szmuda, C. Özdemir, K. Fedorow, S. Ali, and P. Słoniewski, “YouTube as a source of information for narcolepsy: A content-quality and optimization analysis,” *J. Sleep Res.*, vol. 30, no. 2, pp. 1–7, 2021, doi: 10.1111/jsr.13053.
- [2] J. Savigny and A. Purwarianti, “Emotion classification on youtube comments using word embedding,” in *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, Aug. 2017, pp. 1–5, doi: 10.1109/ICAICTA.2017.8090986.
- [3] D. Yasmina, M. Hajar, and A. M. Hassan, “Using YouTube Comments for Text-based Emotion Recognition,” *Procedia Comput. Sci.*, vol. 83, no. Ant, pp. 292–299, 2016, doi: 10.1016/j.procs.2016.04.128.
- [4] P. Y. Saputra, D. H. Subhi, and F. Z. A. Winatama, “Implementasi Sentimen Analisis Komentar Channel Video Pelayanan Pemerintah Di Youtube Menggunakan Algoritma Naïve Bayes,” *J. Inform. Polinema*, vol. 5, no. 4, pp. 209–213, 2019, doi: 10.33795/jip.v5i4.259.
- [5] N. G. Yudiarta, M. Sudarma, and W. G. Ariastina, “Penerapan Metode Clustering Text Mining Untuk Pengelompokan Berita Pada Unstructured Textual Data,” *Maj. Ilm. Teknol. Elektro*, vol. 17, no. 3, p. 339, 2018, doi: 10.24843/mite.2018.v17i03.p06.
- [6] H. Irsyad and M. R. Pribadi, “Implementasi Text Mining Dalam Pengelompokan Data Tweet Pertanian Indonesia Dengan K-Means,” *KURAWAL J. Teknol. Inf. dan Ind.*, vol. 3, no. 2, pp. 164–172, 2020, [Online]. Available: <https://t.co/FXtzMcbdHp>.
- [7] Athifaturrofifah, R. Goejantoro, and D. Yuniarti, “Perbandingan Pengelompokan K-Means dan K-Medoids Pada Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Studi Kasus : Data Titik Panas Di Indonesia Pada 28 April 2018),” *J. EKSPONENSIAL*, vol. 10, no. 2, pp. 143–152, 2019.
- [8] M. M. J. Adnan, M. L. Hemmje, and M. A. Kaufmann, “Social media mining to study social user group by visualizing tweet clusters using Word2Vec,

- PCA and k-means,” *CEUR Workshop Proc.*, vol. 2863, pp. 40–51, 2021.
- [9] W.-L. SONG, “Sentence Similarity Calculating Method Based on Word2Vec and Clustering,” *DEStech Trans. Eng. Technol. Res.*, pp. 374–378, 2020, doi: 10.12783/dtetr/amee2019/33485.
- [10] A. Roihan, P. A. Sunarya, and A. S. Rafika, “Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [11] A. S. R. Sinaga, “Machine Learning Prediksi Karakter Pengguna Hastag (#) Bahasa Generasi Milenial Di Sosial Media,” *Indones. J. Appl. Informatics*, vol. 4, no. 2, p. 165, 2020, doi: 10.20961/ijai.v4i2.41764.
- [12] J. P. Ilmiah, “Text Mining Dan Sentimen Analisis Twitter Pada Gerakan Lgbt,” *Intuisi J. Psikol. Ilm.*, vol. 9, no. 1, pp. 18–25, 2017, doi: 10.15294/intuisi.v9i1.9561.
- [13] S. S. Nagari and L. Inayati, “Implementation of Clustering Using K-Means Method To Determine Nutritional Status,” *J. Biometrika dan Kependud.*, vol. 9, no. 1, p. 62, 2020, doi: 10.20473/jbk.v9i1.2020.62-68.
- [14] D. A. I. C. Dewi and D. A. K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, pp. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.
- [15] P. Jain and S. K. Surve, “Resource Centric Characterization and Classification of Applications Using KMeans for Multicores,” *Int. Conf. Inf. Netw.*, vol. 2019-Janua, pp. 25–30, 2019, doi: 10.1109/ICOIN.2019.8717981.
- [16] A. Nurdin, B. Anggo Seno Aji, A. Bustamin, and Z. Abidin, “Perbandingan Kinerja Word Embedding Word2Vec, Glove, Dan Fasttext Pada Klasifikasi Teks,” *J. Tekno Kompak*, vol. 14, no. 2, p. 74, 2020, doi: 10.33365/jtk.v14i2.732.
- [17] M. Sarifuddin, “Konsep Dasar Makna Dalam Ranah Semantik,” *JISIP (Jurnal Ilmu Sos. dan Pendidikan)*, vol. 5, no. 2, 2021, doi: 10.36312/jisip.v5i2.2024.

- [18] S. Nafinuddin, "Pengantar semantik (pengertian, hakikat, jenis)," in *Pengantar Sematik*, 2020, pp. 1–21.
- [19] S. Ghosh, K. Neha, and Y. Praveen Kumar, "Data wrangling using python," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2 Special Issue 11, pp. 3491–3495, 2019, doi: 10.35940/ijrte.B1427.0982S1119.
- [20] C. R. Harris *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020, doi: 10.1038/s41586-020-2649-2.
- [21] S. Kashikar, S. Patil, A. Vedantwar, S. Katpatal, and S. Pillai, "Weather Prediction using Scikit-Learn," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 4, pp. 36–40, 2019, doi: 10.26438/ijcse/v7i4.3640.
- [22] E. J. Rifano, A. C. Fauzan, A. Makhi, E. Nadya, Z. Nasikin, and F. N. Putra, "Text Summarization Menggunakan Library Natural Language Toolkit (NLTK) Berbasis Pemrograman Python," *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 2, no. 1, pp. 8–17, 2020, doi: 10.28926/ilkomnika.v2i1.32.
- [23] I. A. Ebeid, "Mallet vs GenSim : Topic Modeling Evaluation Report University of Arkansas at Little Rock Information Science , Principles and Theory . Mallet vs GenSim : Topic modeling for 20 news groups report .," no. April 2016, 2019, doi: 10.13140/RG.2.2.19179.39205/1.